

# The seventh art

Moreno-Gelos Quentin

## Introduction

This project enlights the evolution over the years of movie, short and tvSerie making. The adaptation of the way of production as well as our way of consumption are analyzed.

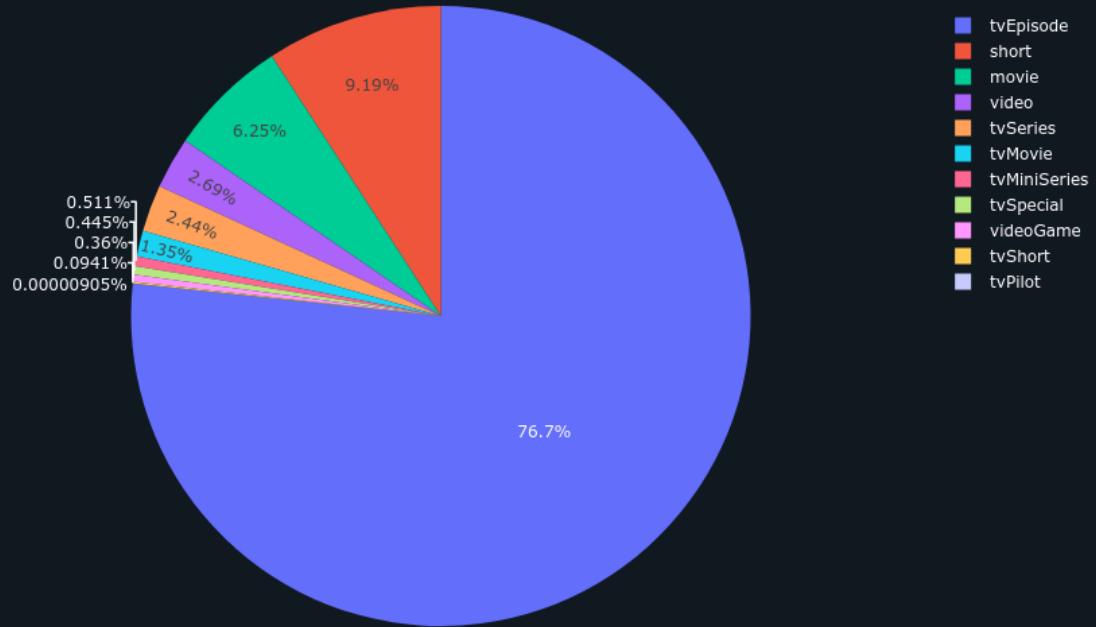


Figure 1: Total production in the database provided by IMDb.

First, let's review the datasets obtained from the IMDb website in October 2024. Figure 1 illustrates the volume of productions included in the datasets. Notably, the majority of the datasets consist of tvEpisode entries, which are the components that make up tvSerie and tvMiniSerie. To gain a clearer understanding of the types of productions made since the early 20th century, we will exclude tvEpisode entries from the dataset, as they are included in the tvSerie and tvMiniSerie categories. Additionally, we will disregard video, videoGame, tvPilot, and tvSpecial entries, as this study focuses specifically on movie (movie, short, tvMovie, tvShort) and series (tvSerie and tvMiniSerie) productions.

The origins of the seventh art can be traced back to March 1895, when Louis and Auguste Lumière held the first screening for approximately 200 members of the Société d'encouragement pour l'industrie nationale in Paris. In December 1895, they hosted a screening for around 40 paying guests and invited acquaintances; this latter event is traditionally regarded as the birth of cinema.

From 1895 to 1915, films were predominantly short productions. One notable example is the famous film *Le Voyage dans la Lune* by Georges Méliès, released in 1902, which lasted only 14 minutes. As shown in Figure 2, long movies began to emerge in 1915, quickly surpassing short films for much of the remainder of the century. By the 1950s, television had replaced radio as the dominant broadcasting medium and became the leading source of home entertainment. Since then, the production of TV programs has steadily increased, as illustrated in Figure 2.

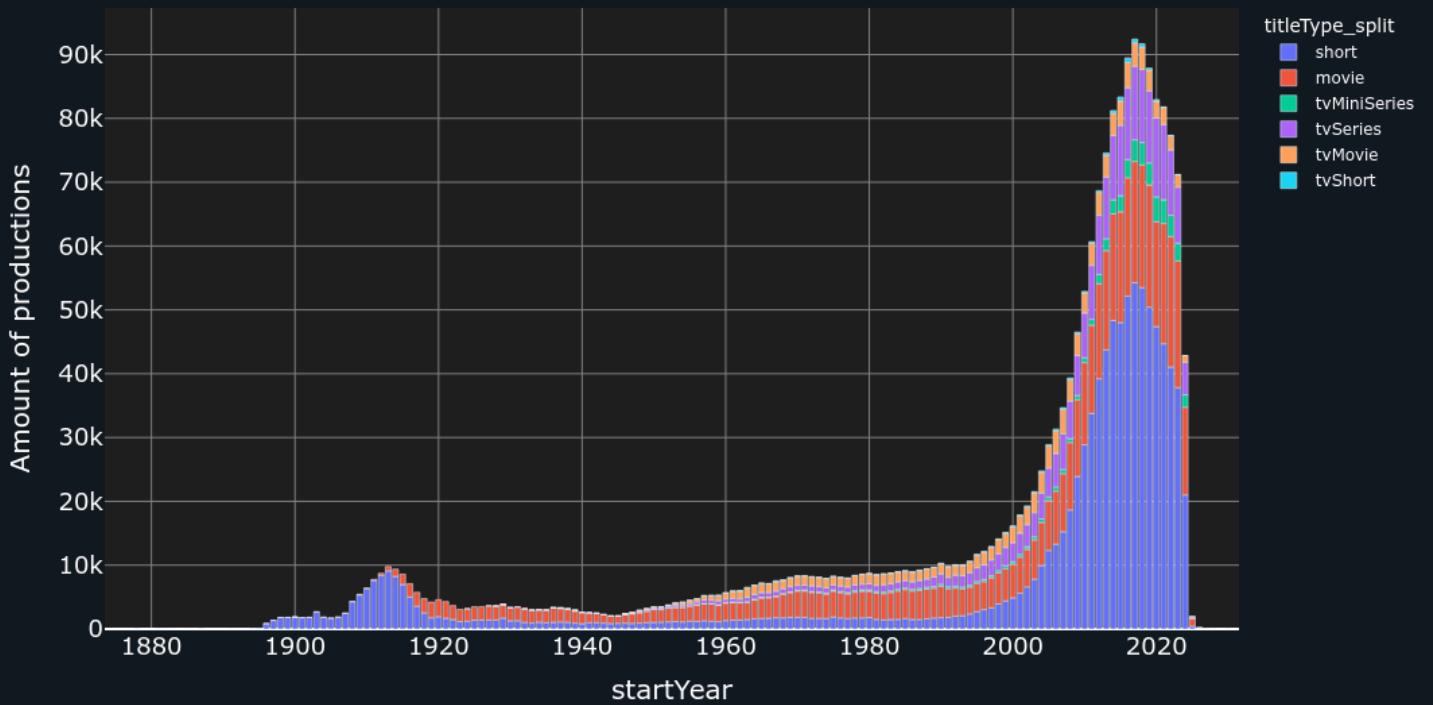


Figure 2: Total production over the years.

Global production has been steadily growing since 1895, with the exception of two periods: 1914-1918 and 1940-1945, coinciding with the two World Wars. Moreover, since the start of the 21st century, production has experienced exponential growth, peaking in 2018.

Since then, certain types of productions have declined significantly, as will be discussed next.

## 1 The movie/short production evolution.

Since 1915, the production of long movies has increased almost linearly and has been the predominant type of production throughout the 20th century, as illustrated in Figure 3. However, this growth has been interrupted by two significant declines over the years. The first occurred during World War II, when production fell from 2,244 to 1,129 movies between 1936 and 1945. The second decline occurred recently, between 2019 and 2020, as a result of the COVID-19 pandemic. Production fell from 19,170 to 16,440 movies within a year. However, in the aftermath of this crisis, the industry rebounded, reaching a record high of 20,443 movies produced in 2022. It is also important to note that tvMovies have become a significant segment of long movie production since 1963, accounting for 35% of the total long movie output. At the beginning of the 21st century, long movie production experienced substantial growth, while tvMovie production remained relatively stable, reaching only 18% of the total in 2016.

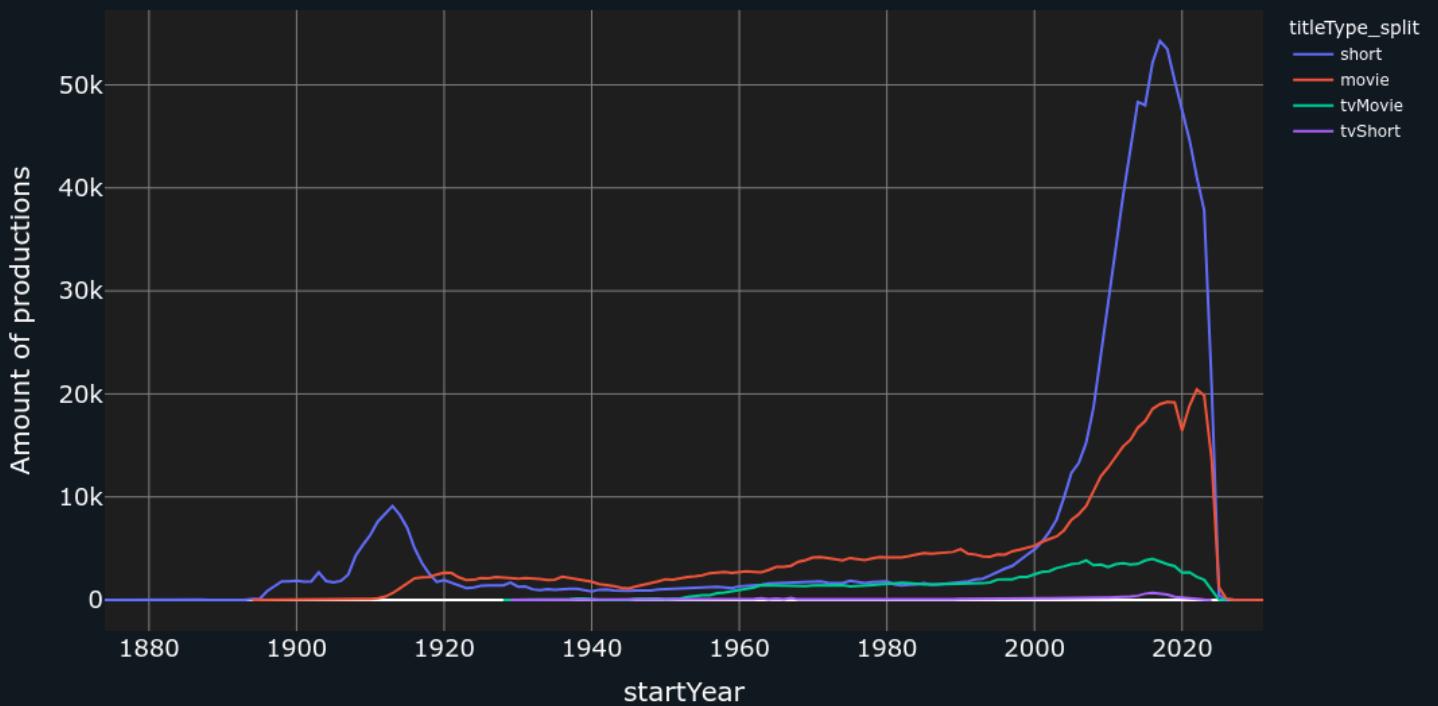


Figure 3: Total movie production over the years proportion short and long movie.

Short movies have existed since 1895, as previously mentioned. The initial rapid growth in the production of short films was interrupted by World War I. While 9,120 shorts were produced in 1913, only 1,757 were made in 1919. After the war, short film production became secondary to long movie production and remained relatively constant for the remainder of the century. Beginning in 1990, an exponential trend in short film production emerged, and shorts once again became the most produced type of film. However, after reaching a peak in 2017 with 54,250 productions, the number of short films has dramatically declined, dropping to 37,820 in 2023—almost half of the total from six years prior. The cause of this decline remains unknown.

## 1.1 The movie production (movie/tvmovie).

There are 28 genres identified for movie production. The most prevalent types of long movies are Drama and Documentary, followed by Comedy and Action. The genre labeled "Unknown" represents incomplete data provided by IMDb, which may introduce uncertainty into our study. The "Other" category includes less common genres, as shown in Figure 4 (Reality-TV, Film-Noir, Talk-Show, Game-Show).

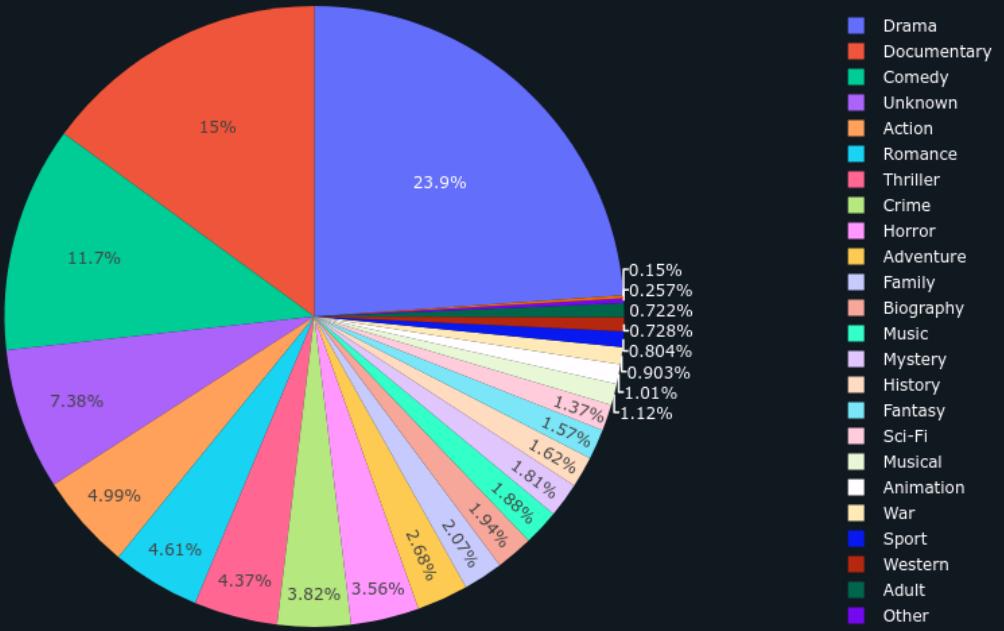


Figure 4: The genres of movie production.

Figure 5 illustrates the evolution of the average runtime, ratings, and number of votes for the main movie genres over the years. The "Other" category encompasses all genres that are not part of the top six productions. Furthermore, to gain a clearer view of the trends in these curves, the data can be smoothed using a Savitzky-Golay filter, as demonstrated in the right panel of Figure 5.

The dynamics of filmmaking have remained consistent across genres over the years. However, a divergence is observed due to the exponential growth rate seen after 1990, as illustrated in Figure 5 (top left). Over more than a century, the length of movies has steadily increased across all genres, following a linear growth trend. In the 1920s, films averaged about 60 minutes, while in the 2020s, they now typically run around 100 minutes, as shown in Figure 5 (second left). The average rating has remained relatively stable at 6/10 over the years, with the exception of documentaries. The popularity of this genre, which has increased linearly since the 1950s, reached a stable value of 7.2 around 1990, as shown in Figure 5 (third from the left). It is important to note that the number of votes for movies produced before 1990 is relatively low, similar to those released in the last few years (see Figure 5, bottom left). This limited voting contributes to significant noise in the data from those early years. Therefore, the accuracy of the ratings should be assessed in relation to the number of votes that informed them.

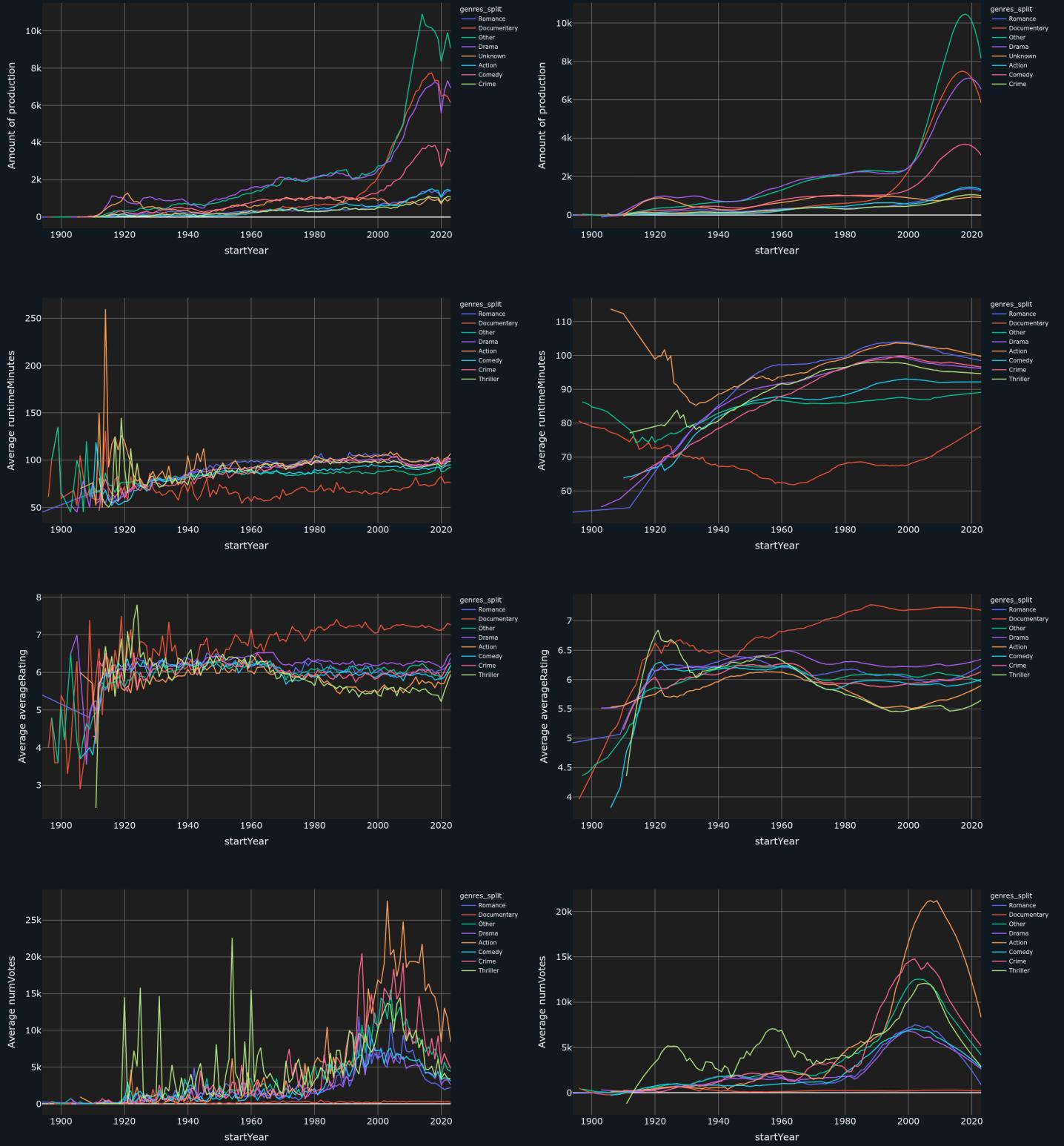


Figure 5: Movies over the years by genre are presented here. The top panel displays total production volume, while the second row shows average runtime in minutes. The third row presents the average rating on a scale of 10, and the bottom panel illustrates the average number of votes. The right panel features data smoothed using a Savitzky-Golay filter of order 1, in contrast to the left panel, which presents the raw, unsmoothed data.

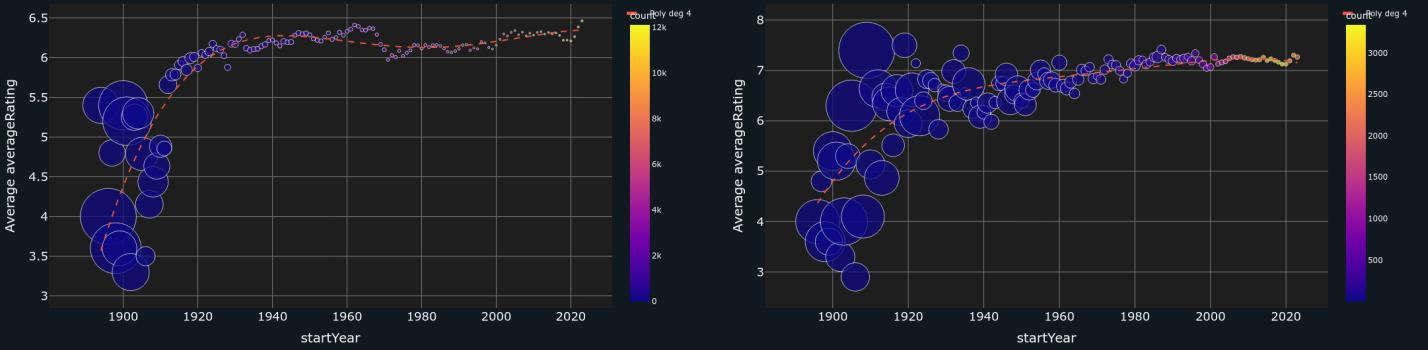


Figure 6: The mean average movie rating over the years is depicted in this analysis. The color bar indicates the number of movies, while the scatter size reflects the associated error, as shown in equation 1. The left panel illustrates total production across all genres, whereas the right panel focuses exclusively on documentary films. Polynomial fits have been applied to both panels, with an order of 4 for the left panel ( $MSE = 0.11$ ) and the right panel ( $MSE = 0.33$ ).

In Figure 6, we present the mean average movie rating over the years. The number of votes has been used to calculate a standard error value, denoted as

$$e_{r,t} = \frac{\frac{1}{N} \sum_{i=1}^N y_{i,t}}{\sqrt{\sum_{i=1}^N z_{i,t}}} \quad (1)$$

which indicates that for each year ( $t$ ), there is an associated error ( $e_{r,t}$ ) related to the mean average movie rating. This error is shown by the size of the scatters. This error is represented by the size of the scatter points. It is clear that the low number of votes for older movies increases the uncertainty in estimating the overall trend.

In the left panel, which displays total production across all genres, we observe that the average rating fluctuates around 6.2, as shown in Figure 5 (third row). Furthermore, movies produced before 1920 appear to be significantly less favored, with a minimum viewer satisfaction rating dipping below 3.5 around 1900. A polynomial fit using machine learning has been applied to both panels of Figure 6. The Mean Squared Error ( $MSE$ ), defined as  $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$  is less than 1 for both fits. Predictions for the upcoming years can be made using the polynomial models. A simple polynomial of degree 3:

$$y = 6.2 - 0.22x + 0.3x^3 \quad (2)$$

is sufficient to describe the trend across all genres, with some adjustments made for documentaries, which exhibit a distinct pattern.

$$y = 6.9 - 0.27x + 0.15x^3 \quad (3)$$

Let us examine the correlation between movie popularity and runtime, as illustrated in Figure 7. It appears that there is a gap in popularity, with the minimum satisfaction level reaching 5.9 out of 10 for movies that are approximately 90 minutes long. It is important to note, however, that this minimum value is not observed across all genres, as shown in Figure 7. Instead, it is a phenomenon that emerges when considering the average across all genres, suggesting that certain genres may have even lower satisfaction ratings. Furthermore, movies longer than 180 minutes and those shorter than 50 minutes are significantly

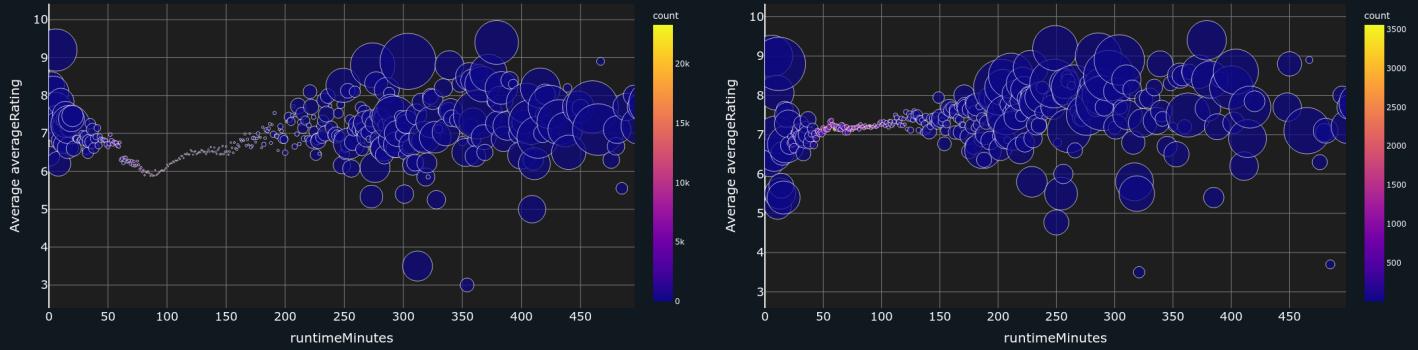


Figure 7: The mean average movie rating over the runtime is depicted in this analysis. The color bar indicates the number of movies, while the scatter size reflects the associated error, as shown in equation 1. The left panel illustrates total production across all genres, whereas the right panel focuses exclusively on documentary films. Polynomial fits have been applied to both panels, with an order of 3 for the left panel ( $MSE = 0.21$ ) and an order of 8 for the right panel ( $MSE = 0.71$ ). Movies with runtimes exceeding 200 minutes have been removed from the dataset.

less common than other films, contributing to the notable divergence in popularity within these runtime extremes.

It is noteworthy that some movies with runtimes below 30 minutes are not categorized as short films in the IMDb database, despite the fact that, as we observed, they represent a negligible fraction of the total.

## 1.2 The short production.

We identified 28 genres in total. The most common types of short films produced are Drama and Comedy, followed closely by Short Documentary and Short Animation. The genre labeled as 'Unknown' reflects compromised data from IMDb, representing an element of uncertainty in our study. Additionally, the 'Other' category encompasses lesser-known genres, as illustrated in Figure 8, including Reality-TV, Film Noir, Talk Show, and Game Show.

The dynamics of filmmaking have remained consistent across genres over the years; however, a divergence has emerged due to the exponential growth rate observed after 1990, as illustrated in Figure 9 (top left). Since the first film screening, the length of movies has steadily increased across all genres over a span of more than 20 years, following a linear growth trend. In the early 1900s, films averaged just a few minutes. After 1920, they did not continue to lengthen significantly, but their average runtime strongly oscillated within the interval of 10 to 25 minutes, as displayed in Figure 9 (second from the left). The average rating shows a clear linear growth over time across all genres, reaching a value between 7 and 7.5, as demonstrated in Figure 9 (third from the left). However it is important to note that the number of votes for short films is extremely low compared to that of longer movies (see Figure 5 and Figure 9, bottom left), resulting in a poor estimation of their ratings. It appears that the audience for this format is inversely related to the age of the short film. We also observe a significant singularity in the horror genre in 1929, with a remarkable 54,293 votes.

In Figure 10, we present the mean average rating of short films over the years. As previously mentioned, short films have gained increasing appreciation, with an improvement of two rating points over the past

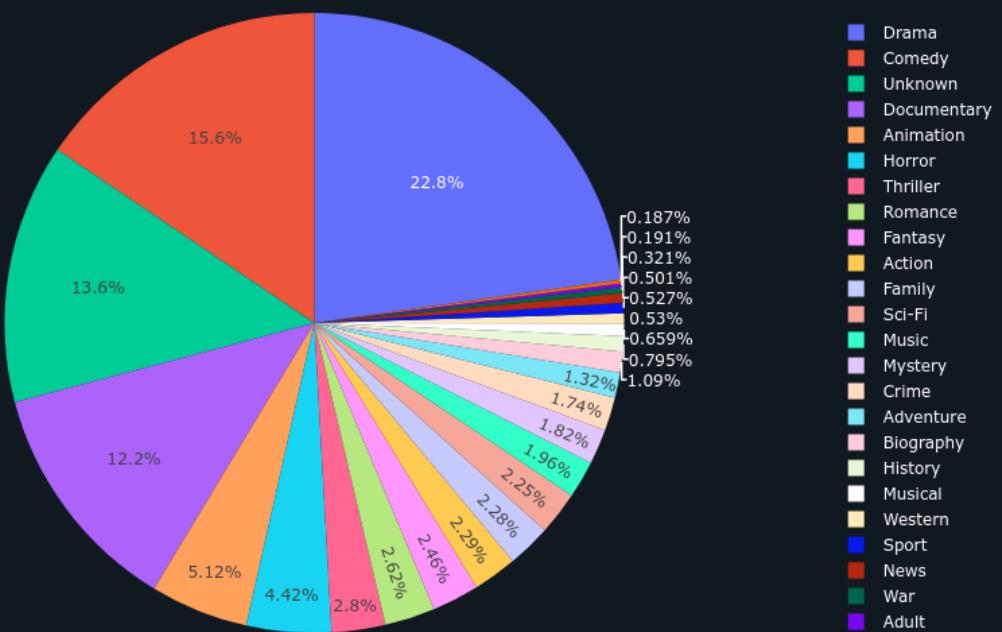


Figure 8: The genres of short production.



Figure 9: Shorts over the years by genre are presented here. The top panel displays total production volume, while the second row shows average runtime in minutes. The third row presents the average rating on a scale of 10, and the bottom panel illustrates the average number of votes. The right panel features data smoothed using a Savitzky-Golay filter of order 3, in contrast to the left panel, which presents the raw, unsmoothed data.

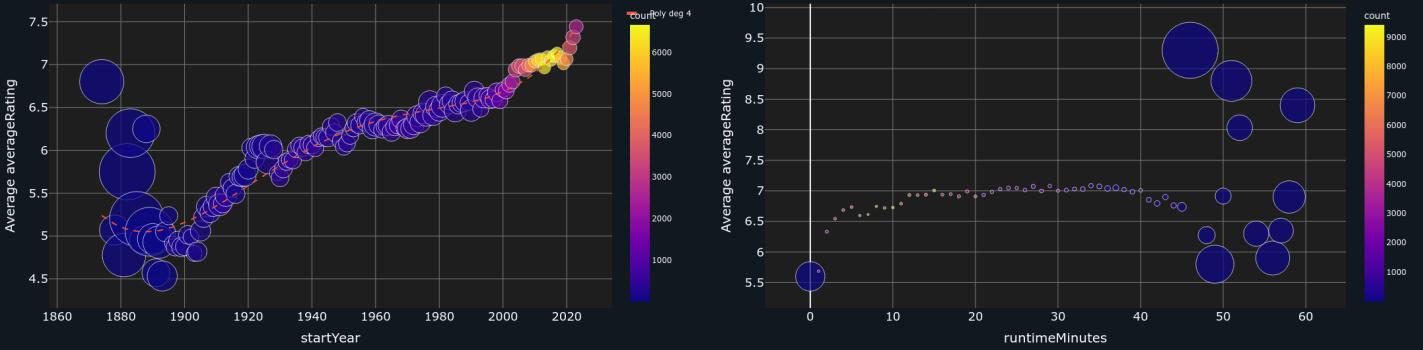


Figure 10: The mean average short rating over the years (left) and the runtime value (right) is depicted in this analysis. The color bar indicates the number of shorts, while the scatter size reflects the associated error, as shown in equation 1. Polynomial fits have been applied to left panel, with an order of 4 ( $MSE = 0.07$ ).

century. This growth can be illustrated by a simple polynomial of degree 4:

$$y = 6.2 - 0.58x - 0.37x^2 + 0.12x^4 \quad (4)$$

It is evident that the short film format does not have the largest audience; however, it surprisingly remains the most produced type of movie today, despite the significant decline in the volume of short productions since 2017. The right panel of Figure 10 shows that once the runtime exceeds 10 minutes, the rating of short productions remains constant at a value of 7.

## 2 The serie production.

In the total series production (TV series, TV mini-series), we identify 27 genres. The most commonly produced series are primarily in the Comedy genre, followed closely by Drama. Additionally, the Documentary and Reality-TV genres are also represented, although the latter genre has been relatively minor in terms of movie production. The genre Unknown represents the compromised data given by IMBD, it can thus be seen as an uncertainty of our study. The 'Other' genre encompasses the lesser-known genres, as illustrated in Figure 11 (including Musical, War, and Western).

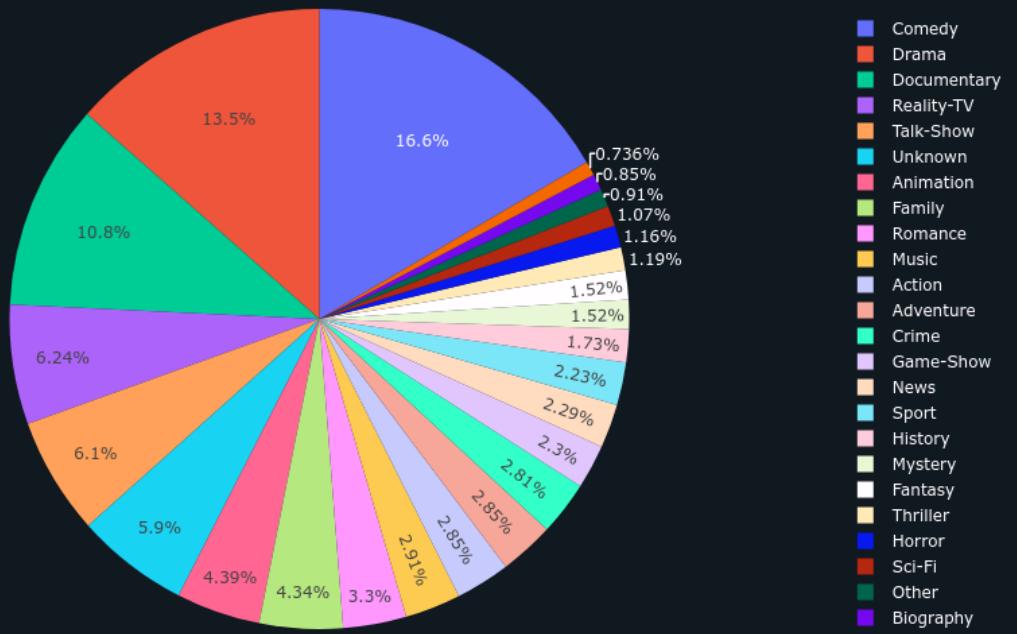


Figure 11: The genres of serie production.

As shown in Figure 12, both types of series emerged in the 1950s and have grown exponentially since then. It is important to note the distinction between TV mini-series and TV series. The distinction between these two types of series lies in that a miniseries is a television show with a predetermined number of episodes that tells a complete story, while a typical TV series often concludes its seasons with cliffhangers and may continue for multiple seasons. In this sense, the miniseries is much more akin to a movie format. It accounted for 24% of the series produced in 2020. Interestingly, while movie production declined in 2020 due to COVID-19, this period marked a peak in series production. However, since then, series production has declined, resulting in a 25% loss in productivity over the three years following COVID-19.

### 2.1 The tvMiniSerie production.

The dynamics of MiniSeries-making have remained consistent across genres over the years. The divergence occurs due to the exponential growth rate observed after 1990, as illustrated in Figure 13 (top left).

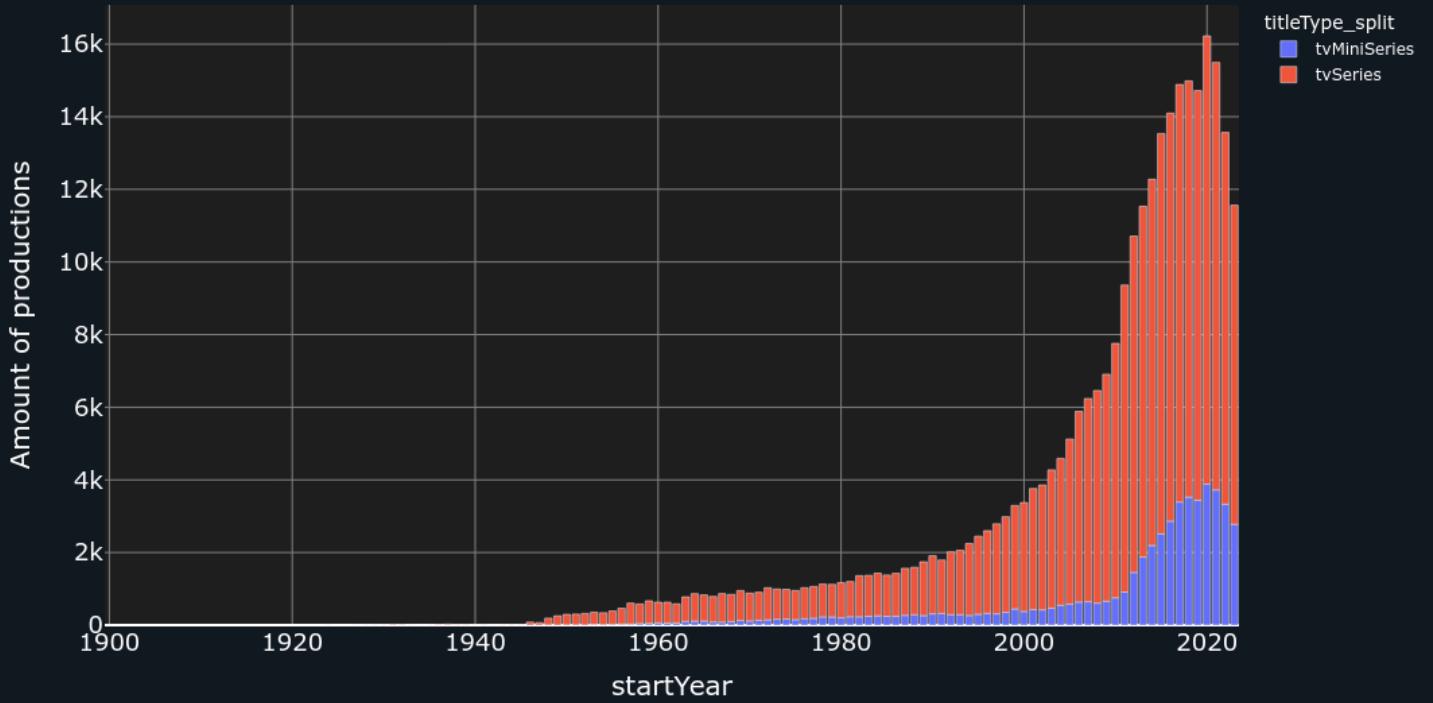


Figure 12: The types of series production.

The length of MiniSeries has steadily decreased across all genres, following a linear trend. In the 1970s, MiniSeries averaged about 150 minutes, whereas in the 2020s, they typically run around 60 minutes, as shown in Figure 13 (second left). The average rating of the top genre trends appears to have slightly decreased over the years; however, recent data indicates an inversion of this trend, with the rating now at a value of 7, as shown in Figure 13 (third left). It is important to note that the number of votes for MiniSeries produced before 1980 is exceptionally low (see Figure 13, bottom left), which explains the significant noise in the data during those early years. Similar to short productions, we observe a significant singularity in the history genre in 2001, marked by a peak of over 30,000 votes.

In Figure 14, we present the average MiniSeries rating over the years and their runtime. In the left panel, which displays total production across all genres, we observe the same tendance saw in Figure 13, with this regain in popularity for MiniSeries made after 2003. It seems that MiniSeries from the 1960s are the most popular, which could explain the significant increase in production after this period.

Figure 14 (right) illustrates the popularity of miniseries in relation to their runtime. It appears that there is a gap in popularity, with the minimum satisfaction level reaching 6.8 out of 10 for MiniSeries that are approximately 170 minutes long. It is important to note, however, that this minimum value is not observed across all genres. Instead, it is a phenomenon that emerges when considering the average across all genres, suggesting that certain genres may have even lower satisfaction ratings.

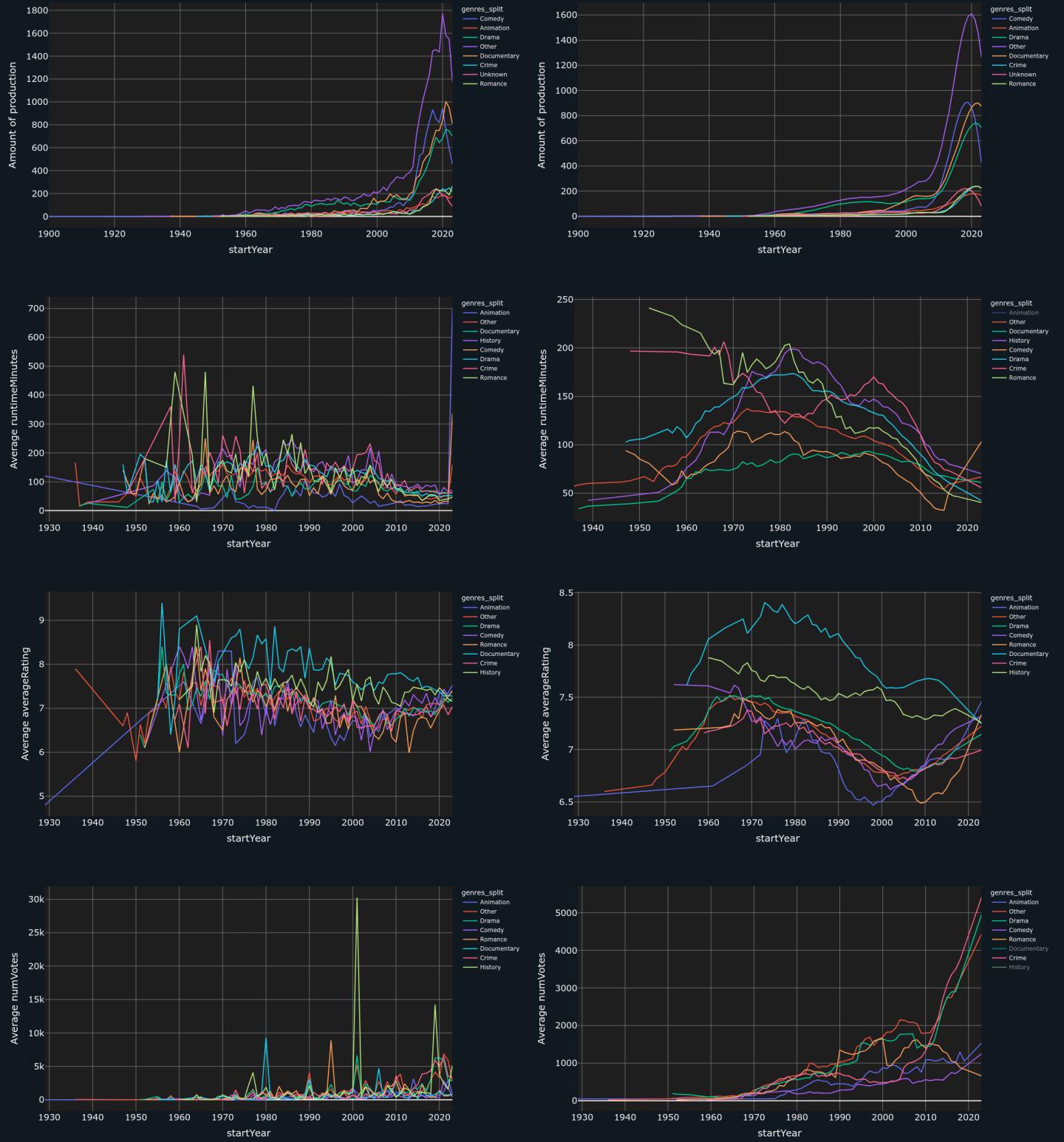


Figure 13: MiniSeries over the years by genre are presented here. The top panel displays total production volume, while the second row shows average runtime in minutes. The third row presents the average rating on a scale of 10, and the bottom panel illustrates the average number of votes. The right panel features data smoothed using a Savitzky-Golay filter of order 1, in contrast to the left panel, which presents the raw, unsmoothed data.

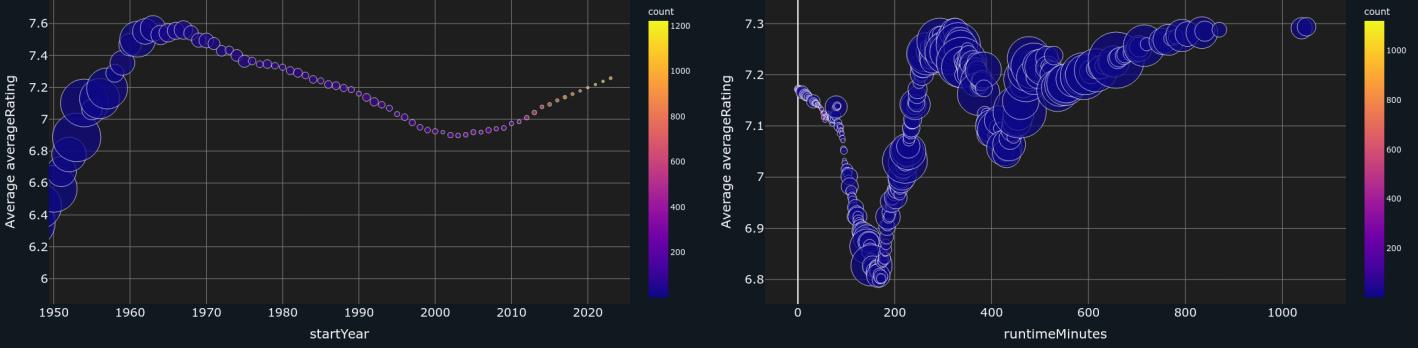


Figure 14: The mean average MiniSerie rating over the years (left) and the runtime value (right) is depicted in this analysis. The color bar indicates the number of MiniSeries, while the scatter size reflects the associated error, as shown in equation 1.

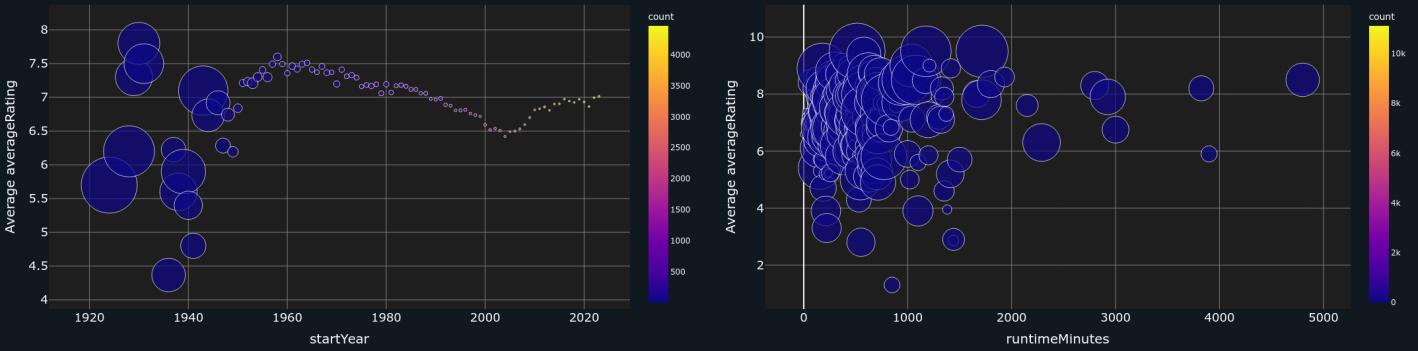


Figure 15: The mean average Serie rating over the years (left) and the runtime value (right) is depicted in this analysis. The color bar indicates the number of MiniSeries, while the scatter size reflects the associated error, as shown in equation 1.

## 2.2 The tvSerie production.

The dynamics of series-making have remained consistent across genres over the years. The divergence arises from the exponential growth rate observed after 1990, as illustrated in Figure 16 (top left). The length of series varies significantly across genres, with notable exceptions such as Reality TV, which experienced a peak runtime of 1,225 minutes in 1982, as shown in Figure 16 (second left). The average rating of the top genre trends seems constant with a value of 7 over the years, as shown in Figure 16 (third left). It is important to note that the number of votes for series produced before 1980 is exceptionally low (see Figure 16, bottom left), which explains the significant noise in the rating data from those early years.

In Figure 15, we present the average Series rating over the years and their runtime. In the left panel, which displays total production across all genres, we observe the same trend seen for the MiniSeries in Figure 14 (left), with this regain in popularity for MiniSeries made after 2003. As for the miniseries, it seems that Series from the 1960s are the most popular, which also explain the significant increase in production after this period. Figure 15 (right) illustrates the popularity of miniseries in relation to their runtime. The divergence is significantly larger when compared to miniseries, making interpretation difficult. Notably,

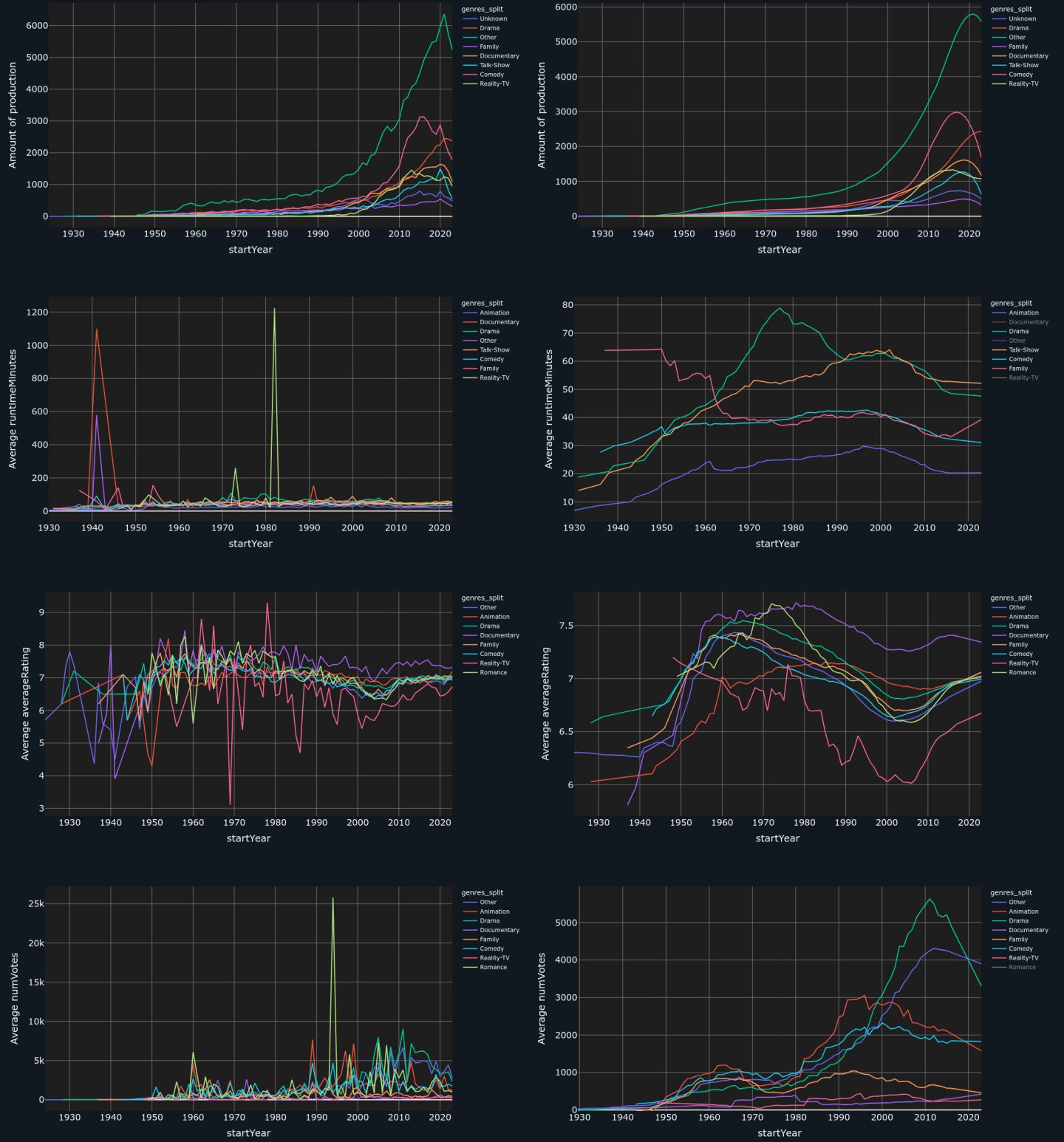


Figure 16: Series over the years by genre are presented here. The top panel displays total production volume, while the second row shows average runtime in minutes. The third row presents the average rating on a scale of 10, and the bottom panel illustrates the average number of votes. The right panel features data smoothed using a Savitzky-Golay filter of order 1, in contrast to the left panel, which presents the raw, unsmoothed data.

some series have extremely long durations, with some reaching almost 5,000 minutes (83 hours).

### 3 The artists expectations

In this section, we will focus on how artists have also evolved over the years in terms of their production methods.

First of all, it is important to note that among the 759,639 directors in the database, 24,503 do not have a recorded birth year. This lack of information, representing approximately 3% of the database, reflects a margin of error in our study. The oldest director in the DataFrame was born in 1818. He is Marius Petipa, the renowned French and Russian ballet dancer, pedagogue, and choreographer. The work attributed to him in the DataFrame is 'Shchelkunchik' from 2023, performed by artists from the legendary Vaganova Academy of Russian Ballet. This production was edited by choreographer Vasily Vainonen and features original multimedia scenery. Although Petipa did not create the movie, his work significantly contributed to its development; therefore, he can rightly be considered one of the directors of this musical film. The youngest director listed is Ali Betil, a producer and director born in 2021, according to IMDb. This exceptionally young age may be attributed to a recent change in his formal name.

Figure 17 (right) illustrates the number of living directors over the years. Since the inception of cinema, the number of directors has steadily increased, reflecting the rise in film production over time (see Figure 2). However, it appears that the number of living directors has reached a maximum of approximately 67,000. This trend is consistent with the decline in the number of first productions since 2012, as illustrated in the left panel of Figure 17.

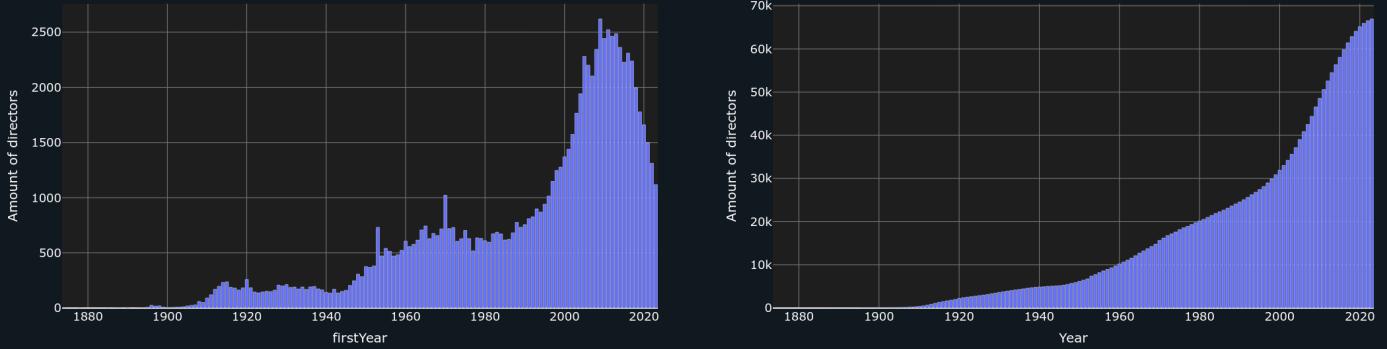


Figure 17: Left panel: Number of directors who made their first production over the years. Right panel: Count of active directors throughout the years.

The age at which directors created their first movie fluctuated around 34 years throughout most of the 20th century, as shown in the left panel of Figure 18. However, if we want to know the mean age of all active directors at the time of their first movie, it appears that this age has gradually decreased from 34.5 years in 1900 to 31.5 years in 2023 (see Figure 18 on the right).

#### 3.1 How much types/genres of production per directors

Figure 19 examines the correlation between the years of directors' first films and the average diversity of genres in their subsequent productions. It is important to note that directors who began their careers after

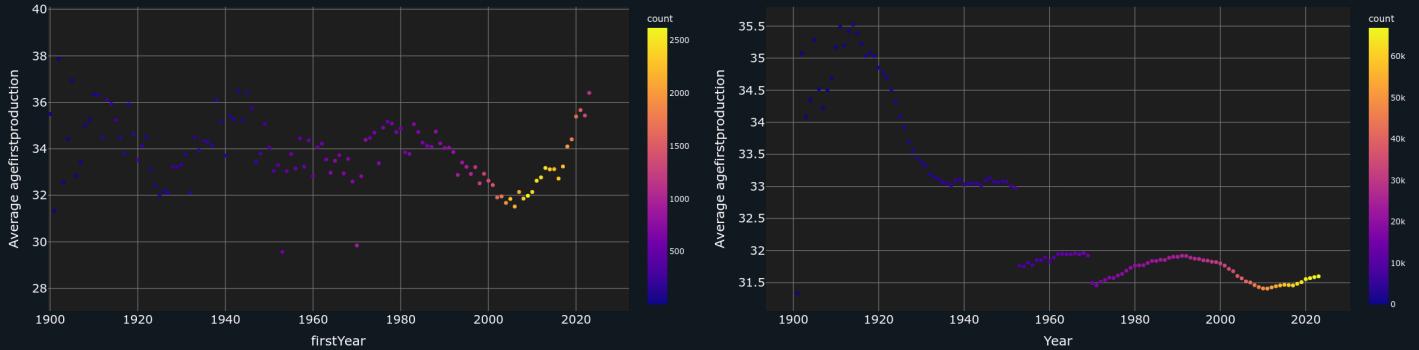


Figure 18: Left panel: Average age of directors when making their first production over the years. Right panel: Average age of all active directors at the time of their first film. The color bar illustrates the number of directors.

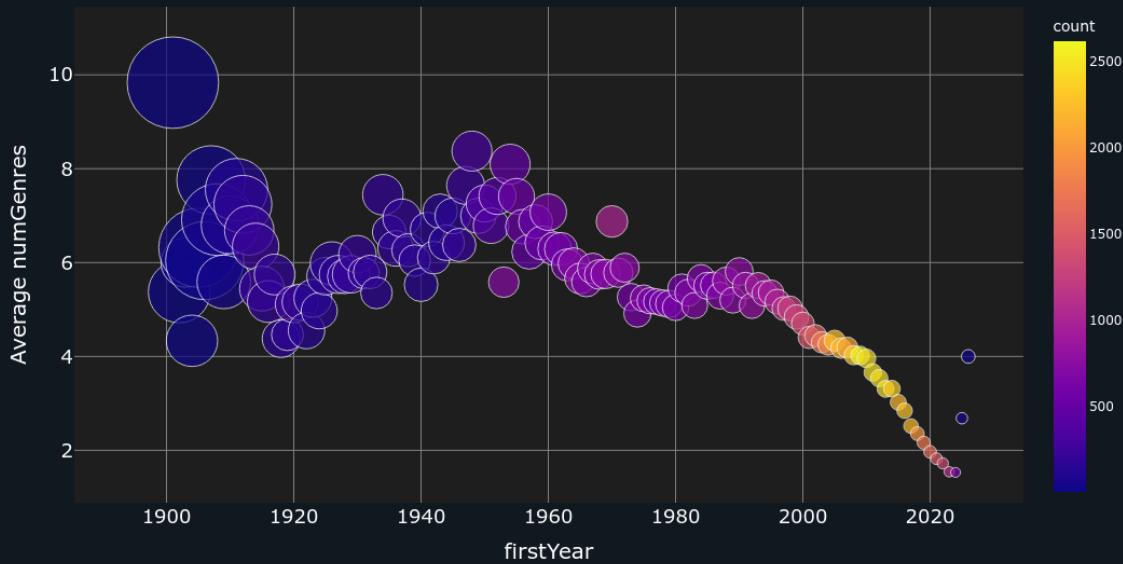


Figure 19: This visualization illustrates the relationship between the years of directors' first films and the average diversity of genres in their subsequent productions. The x-axis represents the years, while the y-axis shows the average number of distinct genres in which the directors have worked. The size of the scatter points indicates the number of productions created by these directors, and the color bar represents the number of directors per year who made their first production.

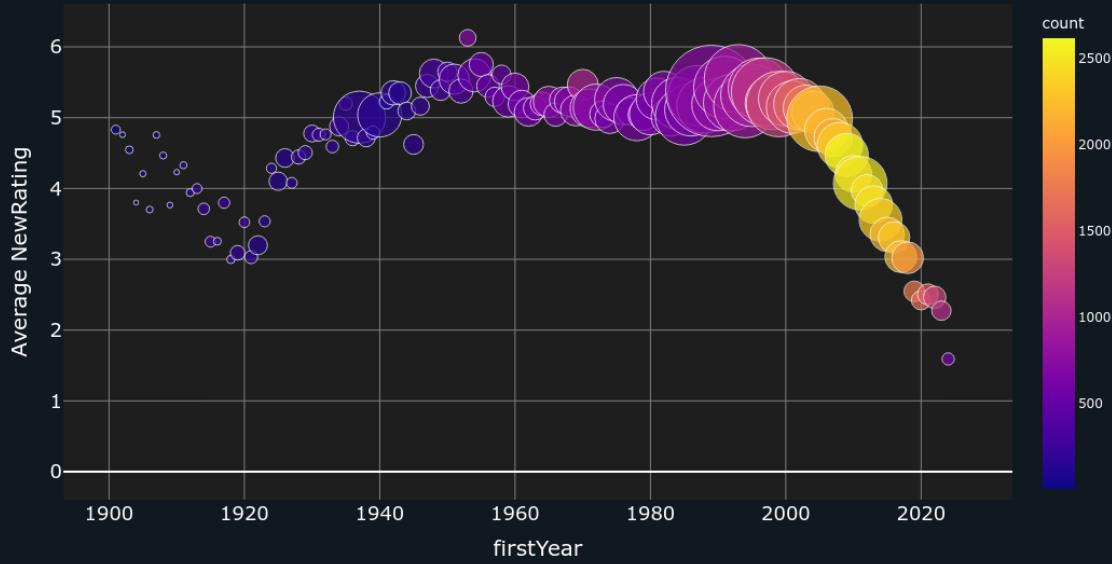


Figure 20: This visualization illustrates the relationship between the years of directors' first films and the average rating in their subsequent productions. The x-axis represents the years, while the y-axis shows the average rating of the production in which the directors have worked. The size of the scatter points indicates the number of votes made for the rating, and the color bar represents the number of directors per year who made their first production.

1980 are more likely to still be active and have the potential to increase their genre diversity in productions. This is also reflected in the declining number of productions per director as the years approach the present (scatter size). This figure can thus be considered nearly accurate only for directors who began their careers before 1980. However, it is evident that both production and genre diversity have decreased over the years, reaching approximately five genres per ten productions around 1980.

### 3.2 How do the ratings of directors evolve over the years?

Figure 20 examines the correlation between the years of directors' first films and the average ratings of their subsequent productions. As noted previously for Figure 19, the interpretation of data after 1980 is somewhat uncertain. It seems that the ratings for older directors, particularly those who began their careers before 1940, are generally low; however, this may be due to a limited number of votes, which introduces some uncertainty. From 1940 to at least 1980, the ratings oscillate around 5, indicating that regardless of the time period, artists consistently find an audience that appreciates their work.

## Conclusion

The exploration of film and series production since the inception of cinema reveals a complex evolution influenced by historical, cultural, and technological changes. Key insights from the various sections of this document indicate a significant shift in the landscape of entertainment, particularly the dominance of series as a format increasingly favored over traditional cinema.

Historically, cinema has transformed from early short films into longer narratives that took center stage in the 20th century. As demonstrated, while the scope of film production has experienced declines—most notably during the World Wars and the recent COVID-19 pandemic—it has nonetheless shown resilience and capacity for renewal. The volume of movie productions remains substantial, with approximately 80,000 movies produced in 2020 compared to around 16,000 series. However, the ratings for series have often surpassed those of films, reflecting a shift in viewer preference toward sustained storytelling and character development that series provide.

This trend indicates that, while cinema continues to dominate in terms of sheer production volume, series have carved out a significant cultural footprint due to their ability to engage audiences deeply over extended periods. The distinction between TV series and mini-series—especially the latter’s rise during peak audience engagement—suggests that serialized content resonates with contemporary viewers’ preferences for immersive narratives.

In examining the genres prevalent in cinema and series-making, it’s evident that while genres like Drama and Comedy dominate, the diversity of content is far broader in series. This trend reflects changing consumer behaviors and preferences for content that spans multiple genres, allowing for experimental storytelling and character development.

Addressing the inquiry of whether series are indeed taking over cinema, it can be observed that although series have not overtaken movies in production numbers, they have increasingly become a compelling option for audiences and critical acclaim. The growing production of series during times when movie production falters—evidenced by the spike during the pandemic—highlights their adaptability within changing entertainment ecosystems.

Looking ahead, the consideration of language and country of origin could add an essential layer to this discussion. Although this analysis did not encompass these aspects due to limitations in the IMDb dataset regarding language and country representations, they are crucial elements in understanding the cinematic landscape. Cinema serves not just as a form of entertainment but also as a lens through which societal values, cultural narratives, and personal experiences are communicated. Exploring how different countries approach film and series production—whether through local language storytelling or cultural context—has the potential to significantly influence viewer engagement and the overall impact of these productions. Countries with robust film industries often reflect their unique cultural expressions in the genres and narratives that dominate their productions, enriching the global cinematic landscape.

## Acknowledgements

The IMDb Non-Commercial Datasets have been used to perform this study. The open source datasets can be found at the following link: [here](#).